Automated Text Mining Comparison of Japanese and USA Multi-Robot Research

Robert J. Watts
U.S. Army Tank Automotive Research, Development & Engineering Center (TARDEC)
http://www.tacom.army.mil/tardec

Alan Porter, Ph.D. and Brian Minsk
Search Technology, Inc
http://www.searchtech.com

Report Documentation Page

Form Approved
OMB No. 0704-0188

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
04 MAY 2004	**Briefing Charts**	**08-01-2004 to 14-04-2004**

4. TITLE AND SUBTITLE
Automated Text Mining Comparison of Japanese and USA Multi-Robot Research

5a. CONTRACT NUMBER
5b. GRANT NUMBER
5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Robert Watts; Alan Porter; Brian Minsk

5d. PROJECT NUMBER
5e. TASK NUMBER
5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Search Technology, Inc,6025 The Corners Parkway ,Suite 202,Norcross,GA,30092

8. PERFORMING ORGANIZATION REPORT NUMBER
; #14060

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army TARDEC, 6501 East Eleven Mile Rd, Warren, Mi, 48397-5000

10. SPONSOR/MONITOR'S ACRONYM(S)
TARDEC

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
#14060

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
Briefing Charts

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**Public Release**	**14**	

Automated Text Mining Comparison of Japanese and USA Multi-Robot Research

SUPERIOR TECHNOLOGY FOR A SUPERIOR ARMY

Briefing Outline:

1. What's Tech OASIS
2. Discuss the Data being Analyzed
   - Field Delimited
   - Multi-Robot Research
3. Processes for Segmenting Data
   - Deductive – Expert Opinion
   - Inductive – PCA based analysis
4. Expectancy Measure
5. Expectancy Measure applied to Segmented data
6. Observations & Interpretations
7. Conclusions & Recommendations

TECH OASIS

vantagepoint

http://www.theVantagePoint.com

Automated Text Mining Comparison of Japanese and USA Multi-Robot Research

Tech OASIS - A Software System for:

- Knowledge Discovery in Large Text Databases
- Profiling Thousands of Research Abstracts

Technology Scanning

Identifying new technologies and new developments in existing technologies

Technology Profiling

Identify key people and organizations

Technology Mapping and Decomposition

Identify dependencies and relationships

Technology Trending

Establish how a technology has emerged

Technology Forecasting

Project how a technology could evolve

Automated Text Mining Comparison of Japanese and USA Multi-Robot Research

FN- DIALOG(R)File 8:Ei Compendex(R)|
CZ- (c) 2004 Elsevier Eng. Info. Inc. All rts. reserv.|
AN- <DIALOG> 06259509|
TI- <MAIN> Guest editorial advances in multirobot systems|
AU- Arai, Tamio^Pagello, Enrico^Parker, Lynne E.|
CS- University of Tokyo Department of Precision Engineering, Tokyo, Japan|
SO- <S2> IEEE Transactions on Robotics and Automation v 18 n 5 October 2002. p 655-661|
DT- JA^(Journal Article)|
AB- <Abstract> As research progresses in distributed robotic systems, more and more aspects of multirobot systems are being explored. This Special Issue on Advances in Multirobot Systems provides a broad sampling of the research that is currently ongoing in the field of distributed mobile robot systems. To help categorize this research, we have identified seven primary research topics within multirobot systems: biological inspirations, communication, architectures, localization/mapping/exploration, object transport and manipulation, motion coordination, and reconfigurable robots. This editorial examines these research areas and discusses the Special Issue papers in this context. We conclude by identifying several additional open research issues in distributed mobile robotic systems. 71 Refs.|
DE- <Descriptors> *Multipurpose robots^Mobile robots^Robotics^Computer simulation|
ID- <Identifiers> Multirobot systems|
CC- <C2> 731.5 _(Robotics)^731.6 _(Robot Applications)^723.5 _(Computer Applications)|

Automated Text Mining Comparison of Japanese and USA Multi-Robot Research

IEEE Imposed Categories "Guest Editorial, Advances in Multirobot Systems"

- 354 EI Compendex & INSPEC abstracts
- Expert Perceived Research Categories
- 324 abstracts grouped

- Deductive Categories
- Expert Field Awareness (e.g., Reconfigurable)

7+2 IEEE Imposed Categories

		IEEE MultiRobot Groups									
	# Records	142	139	114	61	61	44	42	41	30	12
# Records	IEEE MultiRobot Groups	Architecture Allocation Control	Motion Coordination	Communication	Localization Mapping Exploration	Biological	Transport Manipulation Grasping	Robot Learning	Human Interface	IEEE OTHER	Reconfigurable
142	Architecture Allocation Control	142	55	45	27	27	24	14	9		4
139	Motion Coordination	55	139	38	20	22	17	16	19		4
114	Communication	45	38	114	31	25	10	7	12		2
61	Localization Mapping Exploration	27	20	31	61	9	8	5	8		4
61	Biological	27	22	25	9	61	6	5	5		2
44	Transport Manipulation Grasping	24	17	10	8	6	44	7	6		4
42	Robot Learning	14	16	7	5	5	7	42	4		
41	Human Interface	9	19	12	8	5	6	4	41		2
30	IEEE OTHER									30	
12	Reconfigurable	4	4	2	4	2	4		2		12

Automated Text Mining Comparison of Japanese and USA Multi-Robot Research

Tech OASIS Automated Analyses:

- •PCA based factors
- •PCA – NO Singular Factor Solution

PCD Analysis Standardizes PCA

- •Maximizes Inclusion of Abstracts in Factors, Number of Factors & Number of High Loading Factor Defining Terms
- •Minimizes Abstracts in Multiple Factors

Min/Max Analysis - Analogous to Minimizing Entropy & Maximizing Cohesiveness of Factors

2002-03 Multi-robot PCD Factor Groups & Hi-loading Terms

		Descriptors PCD Groups									
# Records	Descriptors	PCD: *OTHER*	PCD: multi-robot systems	PCD: Intelligent robots	PCD: Motion control	PCD: sensor fusion	PCD: multi-agent systems	PCD: Control system analysis	PCD: Robustness (control systems)	PCD: Manipulators	PCD: Collision avoidance
23	multi-agent systems	2					1				
21	multi-robot systems	2	1								
16	cooperative systems	2	1								
14	learning (artificial intelligence)	2					1				
11	Motion control	2			1						
8	Collision avoidance	2									1
7	Control system analysis	2						1			
7	Intelligent robots	2		1							
7	Manipulators	2								1	
7	Robot learning	2					1				
6	Human computer interaction	2						1			
6	sensor fusion	2				1					
5	Robot applications	2			1						
5	Robustness (control systems)	2							1		
5	System stability	2							1		

Expectancy Measure

Likelihood of item in one field having T or more abstracts in a specific category of a second field.

- Cumulative Binomial Distribution
- Detailed View group size / file size defines success probability p
- Field View item frequency *n* times p defines expected frequency
- Cumulative tail calculation based on whether the Detail View item frequency T > or < than expected

If a list item actually occurs *T* times in the records common to the records of a second list item and *T* is greater than or equal to the expected value, we get:

$$p(X \geq T; n, p) = \sum_{r=T}^{n} \left(\binom{n}{r} p^{r} (1-p)^{n-r} \right)$$

Similarly, if *T* is less than or equal to the expected value, we get:

$$p(X \leq T; n, p) = \sum_{r=0}^{T} \left(\binom{n}{r} p^{r} (1-p)^{n-r} \right)$$

Automated Text Mining Comparison of Japanese and USA Multi-Robot Research

SUPERIOR TECHNOLOGY FOR A SUPERIOR ARMY

Expectancy Measure

Likelihood of item in one field having T or more abstracts in a specific category of a second field.

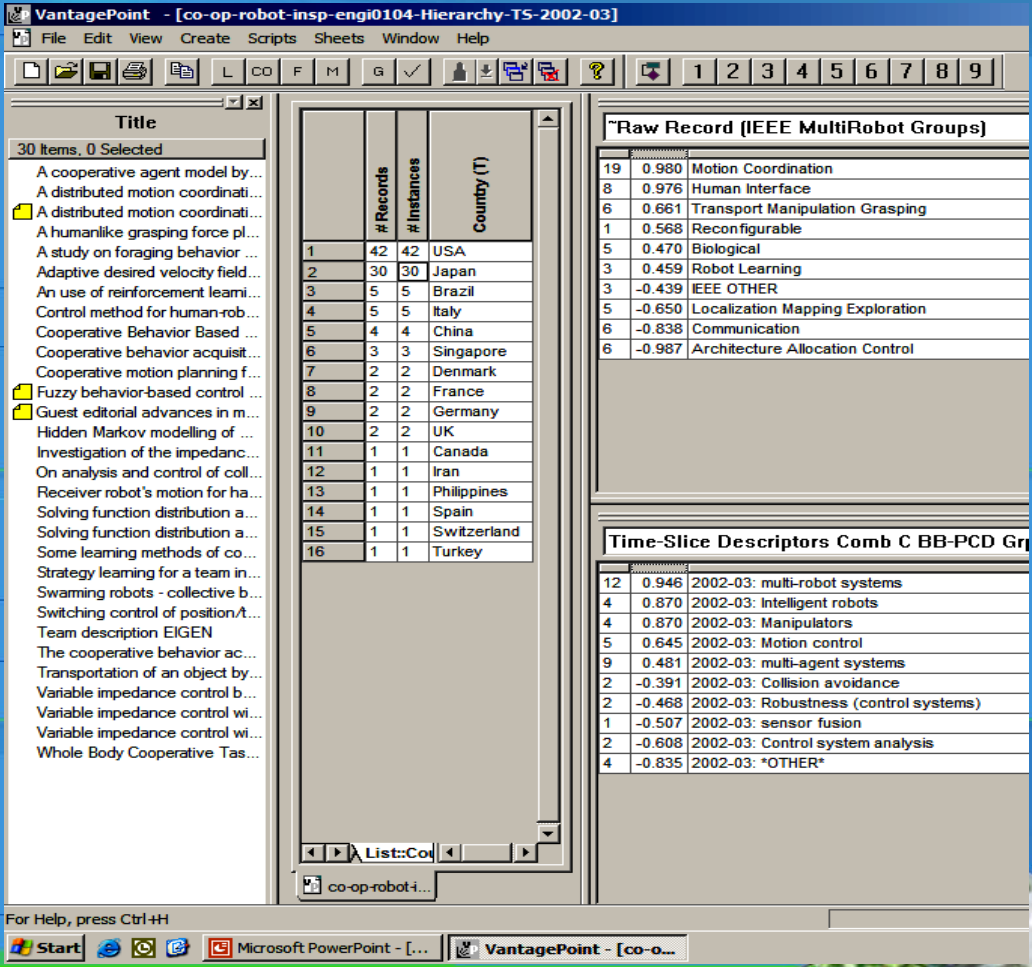
File Size = 107 abstracts
Motion Coordination => 46 abstracts
Probability p = 46/107 = .43

Field View Freq Y = 30
Expected Freq in Detail View = 13

Observed Detail View Freq = 19

Expectancy Measure = 1 -

$$\left[\ p(X \geq T; n, p) = \sum_{r=T}^{n} \left(\binom{n}{r} p^{r} (1-p)^{n-r} \right)\ \right]$$



	# Records	# Instances	Country (T)
1	42	42	USA
2	30	30	Japan
3	5	5	Brazil
4	5	5	Italy
5	4	4	China
6	3	3	Singapore
7	2	2	Denmark
8	2	2	France
9	2	2	Germany
10	2	2	UK
11	1	1	Canada
12	1	1	Iran
13	1	1	Philippines
14	1	1	Spain
15	1	1	Switzerland
16	1	1	Turkey

~Raw Record (IEEE MultiRobot Groups)

19	0.980	Motion Coordination
8	0.976	Human Interface
6	0.661	Transport Manipulation Grasping
1	0.568	Reconfigurable
5	0.470	Biological
3	0.459	Robot Learning
3	-0.439	IEEE OTHER
5	-0.650	Localization Mapping Exploration
6	-0.838	Communication
6	-0.987	Architecture Allocation Control

Time-Slice Descriptors Comb C BB-PCD Gr

12	0.946	2002-03: multi-robot systems
4	0.870	2002-03: Intelligent robots
4	0.870	2002-03: Manipulators
5	0.645	2002-03: Motion control
9	0.481	2002-03: multi-agent systems
2	-0.391	2002-03: Collision avoidance
2	-0.468	2002-03: Robustness (control systems)
1	-0.507	2002-03: sensor fusion
2	-0.608	2002-03: Control system analysis
4	-0.835	2002-03: *OTHER*

TARDEC
U.S. ARMY TANK-AUTOMOTIVE RESEARCH DEVELOPMENT AND ENGINEERING CENTER

Automated Text Mining Comparison of Japanese and USA Multi-Robot Research

SUPERIOR TECHNOLOGY FOR A SUPERIOR ARMY

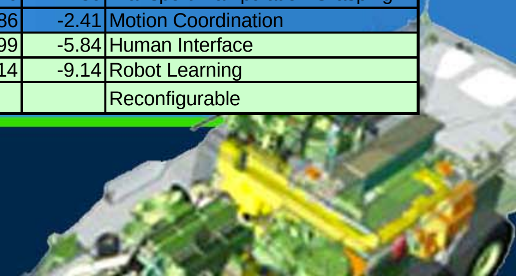
Expectancy Measure
- Anomaly – Expert Input

Low < -0.9
- Protect Competitive Advantage IP
- Publication lull prior to patent applications
- Non-active in area

High > 0.9
- Research Focus Area
- Bias result of National Conference in subject area

IEEE Multi-robot topic areas

Japan Sources					USA Sources				
#	# Grp	Exp.	Metric	Cluster Group	#	# Grp	Exp.	Metric	Cluster Group
1998-99					1998-99				
14	18	0.952	4.28	Human Interface	1	1	0.79	0.79	Reconfigurable
7	9	0.855	3.85	IEEE OTHER	6	18	0.757	1.14	Localization Mapping Exploration
12	18	0.843	2.53	Robot Learning	14	50	0.741	1.03	Architecture Allocation Control
8	13	0.706	1.84	Transport Manipulation Grasping	2	9	-0.356	-1.60	IEEE OTHER
22	44	0.522	1.04	Communication	3	13	-0.366	-1.59	Transport Manipulation Grasping
9	18	0.484	0.97	Localization Mapping Exploration	4	18	-0.42	-1.89	Human Interface
23	50	-0.573	-1.25	Architecture Allocation Control	4	18	-0.42	-1.89	Robot Learning
20	54	-0.926	-2.50	Motion Coordination	12	54	-0.525	-2.36	Motion Coordination
11	34	-0.927	-2.87	Biological	7	34	-0.568	-2.76	Biological
	1			Reconfigurable	9	44	-0.618	-3.02	Communication
2000-01					2000-01				
13	39	0.956	1.43	Motion Coordination	7	9	0.965	4.34	Reconfigurable
4	14	0.641	0.90	Robot Learning	16	38	0.768	1.33	Communication
3	10	0.631	0.90	Human Interface	5	10	0.727	1.45	Human Interface
2	9	0.41	0.53	IEEE OTHER	4	9	0.61	1.10	IEEE OTHER
3	14	-0.359	-1.68	Transport Manipulation Grasping	4	10	0.529	0.88	Biological
7	38	-0.613	-3.33	Communication	7	20	-0.405	-1.16	Localization Mapping Exploration
9	48	-0.638	-3.40	Architecture Allocation Control	4	14	-0.556	-1.95	Robot Learning
2	20	-0.833	-8.33	Localization Mapping Exploration	15	48	-0.668	-2.14	Architecture Allocation Control
	10			Biological	3	14	-0.742	-3.46	Transport Manipulation Grasping
	9			Reconfigurable	10	39	-0.859	-3.35	Motion Coordination
2002-03					2002-03				
19	46	0.98	1.67	Motion Coordination	21	32	0.995	2.89	Communication
8	13	0.976	2.54	Human Interface	20	44	0.759	1.39	Architecture Allocation Control
6	17	0.661	1.02	Transport Manipulation Grasping	11	23	0.718	1.38	Localization Mapping Exploration
1	2	0.568	1.14	Reconfigurable	4	12	-0.515	-1.55	IEEE OTHER
5	17	0.47	0.67	Biological	5	17	-0.676	-2.30	Biological
3	10	0.459	0.66	Robot Learning	5	17	-0.676	-2.30	Transport Manipulation Grasping
3	12	-0.439	-1.76	IEEE OTHER	15	46	-0.786	-2.41	Motion Coordination
5	23	-0.65	-2.99	Localization Mapping Exploration	2	13	-0.899	-5.84	Human Interface
6	32	-0.838	-4.47	Communication	1	10	-0.914	-9.14	Robot Learning
6	44	-0.987	-7.24	Architecture Allocation Control	0	2			Reconfigurable



Automated Text Mining Comparison of Japanese and USA Multi-Robot Research

Expectancy Measure
Anomaly – Expert Input

PCD - fewer hi-low Expectancy grps

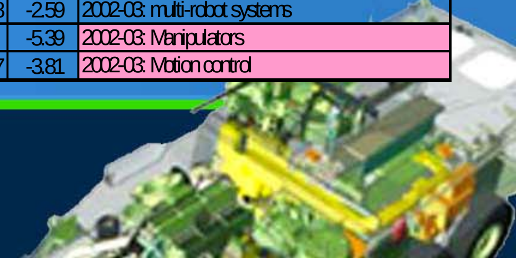
Holistic Approach – Multi-measure Pervasive Findings

PCD Factor Names change over time (e.g., position to motion control and adaptive control to … depicting Tech Maturity

OTHER- Non-consensus

Tech OASIS PCD Groups

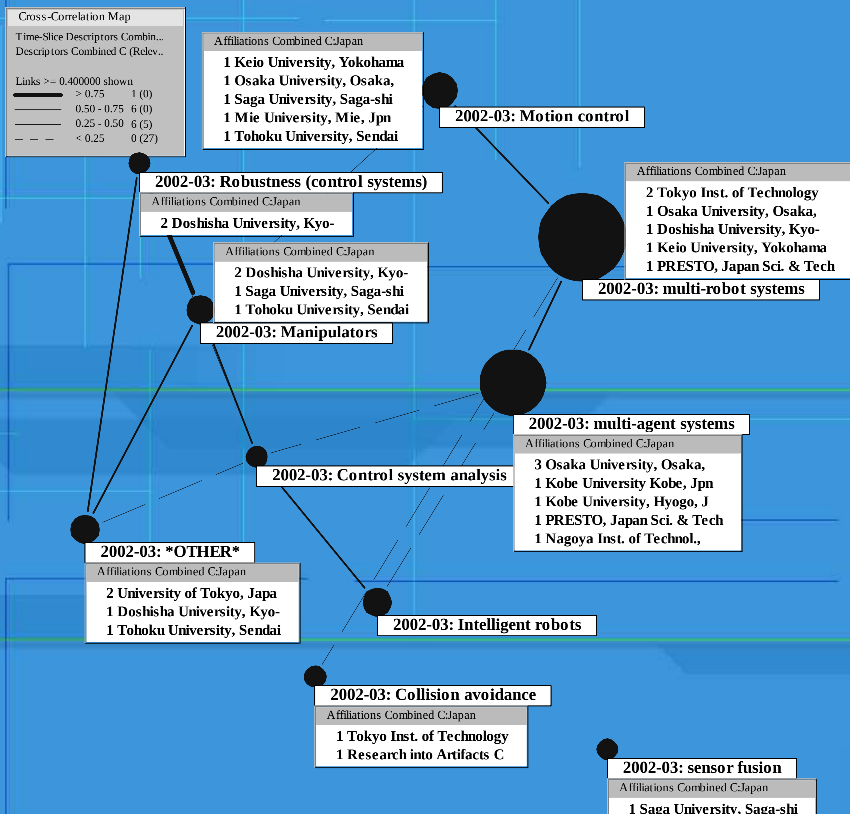
Japan Sources					USA Sources				
#	# Grp	Exp.	Metric	Cluster Group	#	# Grp	Exp.	Metric	Cluster Group
11	16	0.852	2.73	1998-99: Human computer interaction	8	16	0.971	1.94	1998-99: real-time systems
6	10	0.643	1.61	1998-99: Adaptive control systems	5	12	0.853	1.46	1998-99: Robot learning
7	12	0.636	1.53	1998-99: Robot learning	4	10	0.794	1.32	1998-99: Adaptive control systems
6	11	0.554	1.22	1998-99: Position control	6	21	0.628	0.88	1998-99: *OTHER*
10	21	-0.443	-0.93	1998-99: *OTHER*	3	11	0.515	0.71	1998-99: Position control
21	44	-0.478	-1.00	1998-99: learning (artificial intelligence)	4	16	0.471	0.63	1998-99: Human computer interaction
12	26	-0.502	-1.09	1998-99: Intelligent control	6	26	-0.414	-1.79	1998-99: Intelligent control
6	16	-0.671	-1.79	1998-99: real-time systems	6	44	-0.932	-6.83	1998-99: learning (artificial intelligence)
18	43	-0.731	-1.75	1998-99: cooperative systems	5	43	-0.966	-8.31	1998-99: cooperative systems
5	13	0.858	1.39	2000-01: Robot programming	11	24	0.793	1.46	2000-01: *OTHER*
8	25	0.847	1.25	2000-01: Algorithms	6	13	0.696	1.29	2000-01: Robot programming
6	20	0.745	1.06	2000-01: Computer simulation	7	20	-0.405	-1.16	2000-01: Computer simulation
8	34	0.538	0.70	2000-01: multi-robot systems	10	30	-0.495	-1.49	2000-01: multi-agent systems
7	30	0.517	0.67	2000-01: multi-agent systems	3	12	-0.62	-2.48	2000-01: Distributed parameter control systems
3	12	0.511	0.68	2000-01: Distributed parameter control systems	5	19	-0.678	-2.58	2000-01: Manipulators
4	19	-0.394	-1.87	2000-01: Manipulators	7	25	-0.679	-2.43	2000-01: Algorithms
4	24	-0.617	-3.70	2000-01: *OTHER*	9	34	-0.8	-3.02	2000-01: multi-robot systems
12	27	0.946	1.70	2002-03: multi-robot systems	14	24	0.93	2.23	2002-03: *OTHER*
4	7	0.87	2.03	2002-03: Intelligent robots	5	6	0.916	5.50	2002-03: sensor fusion
4	7	0.87	2.03	2002-03: Manipulators	4	11	-0.436	-1.20	2002-03: Control system analysis
5	14	0.645	1.00	2002-03: Motion control	3	9	-0.475	-1.43	2002-03: Robustness (control systems)
9	31	0.481	0.68	2002-03: multi-agent systems	2	7	-0.524	-1.83	2002-03: Intelligent robots
2	8	-0.391	-1.56	2002-03: Collision avoidance	2	8	-0.617	-2.47	2002-03: Collision avoidance
2	9	-0.468	-2.11	2002-03: Robustness (control systems)	10	31	-0.709	-2.20	2002-03: multi-agent systems
1	6	-0.507	-3.04	2002-03: sensor fusion	8	27	-0.768	-2.59	2002-03: multi-robot systems
2	11	-0.608	-3.34	2002-03: Control system analysis	1	7	-0.77	-5.39	2002-03: Manipulators
4	24	-0.835	-5.01	2002-03: *OTHER*	3	14	-0.817	-3.81	2002-03: Motion control



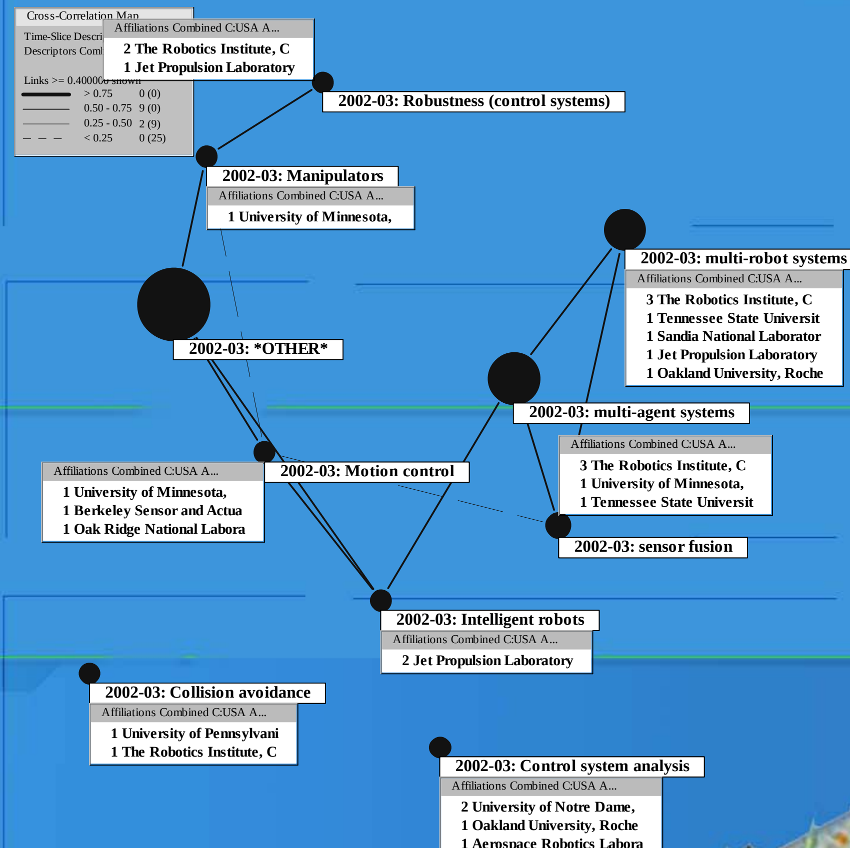
Automated Text Mining Comparison of Japanese and USA Multi-Robot Research

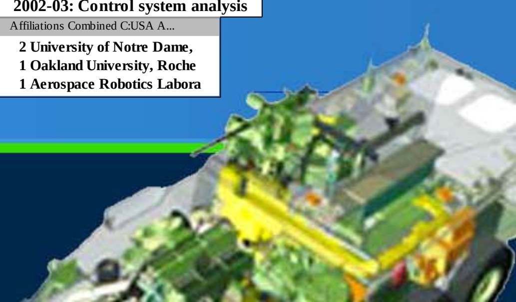
SUPERIOR TECHNOLOGY FOR A SUPERIOR ARMY

Japanese 2002-03 Multi-Robot Research



USA 2002-03 Multi-Robot Research





Automated Text Mining Comparison of Japanese and USA Multi-Robot Research

SUPERIOR TECHNOLOGY FOR A SUPERIOR ARMY

Observations:

Expectancy Measure => Japanese - less emphasis than expected on biological approaches, reconfigurable robots & architecture allocation control (IEEE) and *OTHER* (PCD)
=> Japanese - more emphasis than expected on human interface & motion coordination (IEEE) and Multi-robot systems & manipulators (PCD)

=> USA sources – less emphasis than expected on human interface & robot learning (IEEE)
=> USA more emphasis than expected on reconfigurable robots & communication (IEEE) and sensor fusion & *OTHER* (PCD)

Expert Opinion: Japanese focus more on Industrial Robots and Human Aiding Robots. Must Determine Implications of low *OTHER* expectancy.

Automated Text Mining Comparison of Japanese and USA Multi-Robot Research

SUPERIOR TECHNOLOGY FOR A SUPERIOR ARMY

Conclusions & Recommendations

- Overview of **Tech OASIS & Text-Mining Capabilities**
- Analyzed **Field Delimited Data** on Subject of **Multi-Robot Research**
- Approaches for **Segmentation of the data**:
  - ➢ **Deductive** (i.e., Expert Perceived) **Categories**
    - ✓ Easier to Use to Generalize Observations over time
    - ✓ Field Experts Understand…Acceptance
    - ✓ But…Bias to Present Time Period
  - ➢ **Inductive** (i.e. PCD Derived) **Categories**
    - ✓ Standardizes Analysis
    - ✓ Enables Technology Maturity “Subjective” Assessment
    - ✓ but…Biased by high numbers of low frequency sources of tech papers

Automated Text Mining Comparison of Japanese and USA Multi-Robot Research

SUPERIOR TECHNOLOGY FOR A SUPERIOR ARMY

Expectancy Measure – Ascertain Topical Emphasis Areas & Identifies Unexpected Patterns….as do other measures

Use **Holistic Approach…Pervasive Patterns…Include Field Experts**

Tech OASIS / VantagePoint Automates Clustering / Categorization of Information to Enable and Improve:

- Cognition of Broad Field of Research
- Elicit Research Questions from noted Anomalies
- Promote Innovation through Expert Involvement

TECH OASIS

vantage point

http://www.theVantagePoint.com

TARDEC
U.S. ARMY TANK-AUTOMOTIVE RESEARCH DEVELOPMENT AND ENGINEERING CENTER